



Ripeness Classification of Bananito Fruit (*Musa acuminata*, AA): a Comparison Study of Visible Spectroscopy and Hyperspectral Imaging

Yuan-Yuan Pu¹ · Da-Wen Sun¹ · Marina Buccheri² · Maurizio Grassi² · Tiziana M.P. Cattaneo² · Aoife Gowen¹

Received: 8 January 2019 / Accepted: 9 April 2019 / Published online: 3 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Bananito (*Musa acuminata*, AA) fruit in three maturity stages (stages 2, 4 and 6) were investigated in this study. Correlation analysis was conducted between fruit firmness, soluble solids content (SSC) and several colour parameters. Point-measured spectroscopy (Vis-point) and spatial-measured hyperspectral imaging (Vis-HSI) were applied to collect visible spectra (400–740 nm) from the fruit peel. Three classification methods, k-nearest neighbour (k-NN), soft independence modelling of class analogy (SIMCA) and partial least square discriminate analysis (PLSDA), were applied for maturity stage classification. Results showed that a strong correlation was found between SSC and peel yellowness index ($r = 0.92$). Ripeness classification models developed using Vis-HSI data performed better than using Vis-point data. The best model based on PLSDA achieved a total correct classification rate of 93.3%. A simplified PLSDA model established on three wavelengths (650 nm, 705 nm and 740 nm) derived from the regression vector provided an equivalent model performance. This study demonstrated the use of hyperspectral imaging for accurate and non-destructive ripeness classification of bananito fruit based on the visible peel spectra, and the potential of using the three feature wavelengths to develop a multispectral imaging system for industrial application.

Keywords Hyperspectral imaging · Spectrometer · Visible · Bananitos · Maturity · Classification · PLSDA

Introduction

Banana is generally cultivated in tropical and subtropical areas in over 130 countries around the world (Mohapatra et al. 2011). It is a popular fruit in people's "five a day" menu in many European countries, where the consumption of at least five servings of fruits and vegetables per day is encouraged. Bananito (*Musa acuminata*, AA), also called baby-banana or mini-banana, has a smaller fruit size compared to normal bananas. It provides all nutrients that the conventional bananas have, but in an intense, sweeter and creamy flavour (Li et al. 2006). In addition, the bananito fruit is easy to digest (FreshPlaza 2018) which is a distinction to normal bananas.

As a result, the exotic bananito fruits are very welcomed by the people who live outside the cultivated regions.

The ripeness level of banana fruit is vitally important for fruit transportation and storage, quality control and marketability (Li et al. 2011). Conventionally, bananas maturity stages are mainly determined by the following methods: (1) visually evaluating the whole-fruit colour and comparing the colour with a standard colour chart (Hashim et al. 2012; Pathare et al. 2013); (2) measuring fruit firmness using a penetrometer; (3) or measuring some components such as the pulp to peel ratio, soluble solid content, titratable acidity and starch pattern (Robinson and Saúco 2010). These strategies are based on the evaluation of the apparent colour, texture or other chemical properties that are associated with fruit ripening. However, the determination procedure is subjective (visual evaluation), or time-consuming and sample-destructive, which is not practical for widespread industrial application.

In recent few decades, several non-invasive techniques have been applied to bananas' quality attribute prediction, maturity stage classification and disease identification. For example, the use of colourimeters for monitoring colour changes during storage (Kajuna et al. 1998) and predicting textural attributes of bananas (Jaiswal et al. 2014), the use of

✉ Da-Wen Sun
dawen.sun@ucd.ie; <http://www.ucd.ie/refrig>; <http://www.ucd.ie/sun>

¹ School of Biosystems and Food Engineering, University College Dublin, Belfield, Dublin 4, Ireland

² Council for Agricultural Research and Economics (CREA-IT), Via Venezian 26, 20133 Milan, Italy

electronic nose (Eduard et al. 1999) or gas sensors (Steffens et al. 2010) to analyse the volatile components emitted from banana fruits for ripeness stages determination and the use of RGB images captured by a computer vision system (Mendoza and Aguilera 2004; Sanaeifar et al. 2016; Surya Prabha and Satheesh Kumar 2015) for prediction of bananas' maturity stages and some quality indices (i.e. firmness, total soluble solids, pH). Chilling injury occurred on bananas could be inspected by laser-induced backscattering (Hashim et al. 2013) and senescent spotting symptoms can be detected by the use of fractal texture analysis on banana images (Quevedo et al. 2008). Due to physicochemical changes taking place in the fruit during ripening that will bring variations to corresponding spectra, the spectroscopic technology is useful in the determination of banana quality attributes and maturity stages (Zude 2003; Davey et al. 2009; Liew and Lau 2012).

Among the above-mentioned non-destructive methods, the spectroscopic technique gained great attention in industry because the instrument is easy-operating and cost-effective, with capabilities in fast and accurate determination of a variety of agri-food products (Rambo et al. 2016; Liu et al. 2011; Magwaza et al. 2012; Cortés et al. 2017, 2019). In recent decades, hyperspectral imaging (HSI) technique has emerged as an attractive non-contact process analytical tool for agri-food quality evaluation (Wang et al. 2015a; Pu et al. 2015; Liu et al. 2017, 2018; Ma et al. 2017, 2018; Pan et al. 2018; Cheng et al. 2016a, 2016b, 2017, 2018; Dai et al. 2016), as the combination of computer vision (Du and Sun 2005; Jackman et al. 2011; Sun and Brosnan 2003; Wang and Sun 2003; Zheng et al. 2006) and spectroscopy (Morsy and Sun 2013; He and Sun 2015; Wang et al. 2017a, 2017b; Xu et al. 2015) in the HSI system provides spectral and spatial information of target objects in one scan. For banana fruit, Rajkumar et al. (2012) proposed the use of line-scanned hyperspectral imaging in visible and NIR regions (400–1000 nm) for some quality attributes prediction and maturity stages determination. Multiple linear regression models developed on the selected wavelengths yielded good performance ($R^2 > 0.85$) for firmness, moisture content and total soluble solids prediction. Correlations between each quality attribute and the maturity stage at different temperature have also been investigated. Wang et al. (2015b) studied the potential of using reflectance hyperspectral imaging (400–1100 nm) for banana shelf-life prediction in terms of different browning levels. By combining image features extracted from principal component analysis (PCA) and the average spectra, a classification model based on back-propagation neural network algorithm achieved the best performance for shelf-life prediction.

The distinction between bananito fruit and banana fruit is obvious in terms of texture (i.e. firmness) and sweetness (i.e. soluble solids content). Though many investigations have reported the quality changes of conventional banana fruit during ripening process, as well as the application of spectroscopic

and imaging techniques for maturity stages determination, there is a lack of knowledge in understanding the quality evolution of bananito fruit (*Musa acuminata*, AA) during ripening. Furthermore, a comparison of the performance between spectroscopy and HSI techniques on bananitos ripeness prediction has never been made. The current study applied hand-held spectrometer and desktop hyperspectral imaging to capture visible spectral information from bananito fruits, with the objectives of (1) understanding the changes of fruit colour, soluble solids content and firmness of bananito fruit when it ripens, (2) investigating the correlation between fruit firmness, soluble solids content and several colour parameters, (3) comparing classification performance of spectrometer and hyperspectral imaging for bananitos ripeness stage.

Materials and Methods

Fruit Samples

Bananito fruit (*Musa acuminata* (AA), c.v. Pisang mas, originates from Colombia) at maturity stage 2, stage 4 and stage 6 were kindly supplied by a fruit company (AL.MA s.r.l.) in Milan, Italy. The maturity stages were defined by visually inspecting the whole-fruit skin colour, using a bananito ripeness chart provided by the fruit supplier as a reference. This classification is similar to the ripeness chart that is commercially used for banana fruit (USDA 2001). The weight (whole-fruit) and diameter (measured in the central part of the fruit) of all bananito fruits were determined, and results are presented in Table 1.

Instruments and Spectra Acquisition

A portable spectrophotometer (CM-2600d, Konica Minolta, Inc., Japan) covering spectral region of 360–740 nm was used to collect the point-measured reflectance spectra from bananitos peel. This instrument contains a diffused illumination system (D65, 2° standard observer) with an integrating sphere size of 52 mm in diameter. The spectral separation device is diffraction grating, and the measuring window is 6 mm in diameter. The spectrophotometer measurements were carried out on three different spots (stalk-end, middle part and

Table 1 Details of bananito samples in three maturity groups

Ripeness stages	2	4	6
No. of fruits	30	26	34
Diameter (mm)	22.51 ± 1.19	22.57 ± 1.21	23.03 ± 1.48
Weight (g)	47.00 ± 6.26	48.90 ± 5.31	48.64 ± 7.07

The diameter and weight data are presented as “mean ± standard deviation” for each ripeness stage

tip-end) on each side of a fruit. As a result, a total of six measurements per fruit were collected.

A desktop hyperspectral imaging system (DV S.r.l., Padova, Italy) was employed to simultaneously acquire spatial and spectral data from bananitos peel. The HSI system mainly consists of a CCD camera (avA 1000-100gm, Basler AG, Germany), a spectrograph (V10H, Spectral Imaging Ltd., Finland) providing spectral information from 400 to 1000 nm with a resolution of 5 nm and a cylindrical diffuser equipped with 150 W halogen lights. Hyperspectral images were taken in a dark room to avoid interference from external lights. Both sides of whole-fruits were scanned. The two instruments were carefully calibrated before data acquisition. The spectrometer is calibrated using a white calibration plate provided by the supplier. The HSI system is calibrated by a white and a dark reference image as detailed in the paper of Gowen et al. (2009).

Colour Measurement

Spectral data of bananitos peel was analysed by the spectrophotometer and transformed in L^* , a^* , b^* colour parameters (Mclaren 1976). L^* represents lightness and ranges from 0 (black) to 100 (white). The chromaticity coordinate a^* gives the range of green ($-a^*$) and red ($+a^*$) colour, while b^* indicates the blue ($-b^*$) and yellow ($+b^*$) colour. *Chroma* (saturation) and hue angle (*hue*, in radians) can be calculated as:

$$\text{Chroma} = \sqrt{(a^*)^2 + (b^*)^2} \quad (1)$$

$$\text{hue} = \tan^{-1}(b^*/a^*) \quad (2)$$

The calculation result of *hue* from Eq. (2) was transformed to degrees (h°) and used in the study. By transforming $L^*a^*b^*$ colour into XYZ colorimetric space, yellowness index (*YI*) (ASTM 1988) of bananito peel can be calculated using Eq. (3).

$$YI = [100(1.28X - 1.06Z)]/Y \quad (3)$$

Firmness and SSC Measurement

Two internal quality indices (firmness and soluble solid content) of each sample were determined in the study. Firmness measurement was conducted using an Instron Universal Testing Machine (Model 4301, Instron Ltd., UK, software Bluehill 2.35) equipped with a cylindrical flat test probe (\varnothing 6.5 mm, cross head speed 200 mm/min). After peel removal, flesh firmness was measured on the middle part of the fruit. The testing machine evaluated the maximum force needed to penetrate the fruit to 8 mm. Bananito fruits were then cut crosswise in two halves and frozen. SSC analysis was carried out on each half: 5 g of frozen flesh was diluted with milliQ water (1:3 w/w). The mixture was blended for 30 s (T10 Ultra-Turrax®, IKA, DE) and then centrifuged (6000×g, 4 °C, 10 s) (Blankenship et al.

1993). A drop of the supernatant was assessed for SSC (two readings per each bananito half) by means of a refractometer (RFM 81, Bellingham+FFvalues (%)) were averaged per fruit.

Data Processing

Statistical Analysis

Statistical analysis including analysis of variance (ANOVA), post hoc analysis based on Bonferroni test and correlation analysis between firmness, SSC and all the colour parameters were conducted in Statgraphics 5.1 software (Statpoint Technologies, Inc., USA).

Extracting Spectrometer and HSI Data

Mean Spectrum Extraction For the spectrometer data, a mean-spectrum was used to represent each bananito sample by averaging all spectra collected from the six different locations. For the HSI data, the process of extracting mean-spectrum for each bananito whole-fruit is shown in Fig. 1. After hyperspectral image acquisition, a region of interest (ROI) was selected by removing both ends (tip-end and stalk-end) from the whole-fruit. Background removal was performed on the image at wavelength 695 nm using simple thresholding method (Pu and Sun 2016). The mean-spectrum was then calculated by averaging all pixel-spectra that belonged to the sample. Since the two sides of a whole-fruit were scanned, the two mean-spectrum obtained from both sides were averaged to get a final mean-spectrum that can represent a whole-fruit sample.

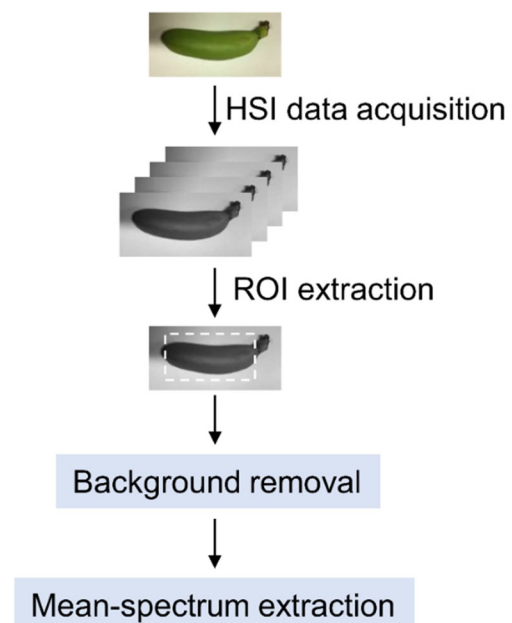


Fig. 1 Extraction of mean-spectra from hyperspectral imaging data

Trim Spectral Region and Interpolation The spectral region and resolution covered by the spectrometer (range, 360–740 nm; resolution, 10 nm; no. of wavelengths, 39) and the hyperspectral imaging system (range, 400–1000 nm; resolution, 5 nm; no. of wavelengths, 121) were different. To compare classification and prediction results in a fair way, we trimmed the spectral regions into the same range. In the study, the 400–740 nm region was covered by the two instruments, thus, this region was considered as a common spectral region. The spectra obtained by the spectrometer were interpolated to have the same spectral resolution as the HSI system (5 nm). Thus, the spectra from both instruments were in consistency, with spectral range of 400–740 nm, spectral resolution of 5 nm, and wavelengths numbers of 69. Results of spectral trimming and interpolation are shown in Fig. 2 a and b.

Data Set Splitting The ninety bananito fruits at three different maturity stages (stages 2, 4 and 6) were divided into calibration set (two-third of the total fruits) and validation set (one-third of the total fruits). To make sure that samples in calibration and validation set were well represented (both datasets had a similar proportion of fruits in three maturity stages), all fruits were firstly labelled and put in order, namely fruit no. 1–30 were in stage 2, fruit no. 31–56 were in stage 4 and fruit no. 57–90 were in stage 6. Then, the fruits with a number of $3*n$ ($n = 1, 2, \dots, 30$) were selected into validation set and the remaining fruits were in calibration set. Thus, in the calibration set, there were 20, 17 and 23 fruits in maturity stage 2, 4 and 6, respectively, and in the validation set, there were 10, 9 and 11 fruits in maturity stages 2, 4 and 6, respectively.

Multivariate Data Analysis

Spectral pre-processing and multivariate modelling were performed using PLS_toolbox 8.2 (Eigenvector Research, Inc., USA) and Matlab 2015a (The Math Works, Inc. USA).

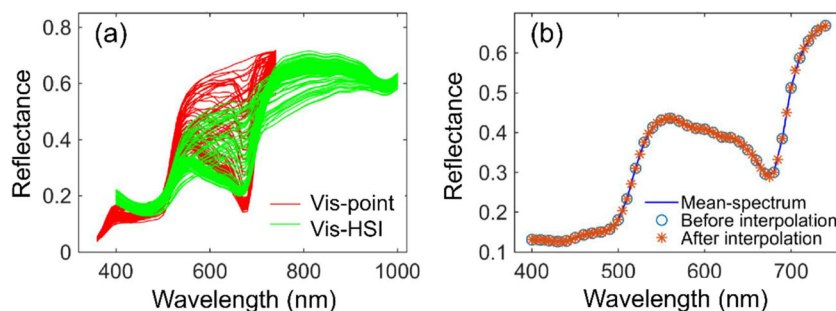
Spectral Pre-processing The obtained spectra after trimming and interpolation were pre-treated with standard normal variate (SNV) in combination with mean-centring before any multivariate models were built. The selection of the pre-processing

approaches was based on a preliminary investigation by trying different pre-processing methods to build a classification model.

Principal Component Analysis As a preliminary analytical approach, principal component analysis (PCA) was applied to 90 samples in order to investigate their relationships with different maturity stages. PCA finds orthogonal variables called principal components (PCs) to explain data variance. Each PC is a linear combination of the original variables. Since the first few PCs capture the most significant information in the data, PCA has been proved to be an effective method for data exploration and data reduction in spectroscopic analysis (Ravikanth et al. 2017).

Ripeness Classification Three supervised classification methods, namely soft independent modelling by class analogy (SIMCA), partial least square discrimination analysis (PLSDA) and k-nearest neighbour (k-NN) were applied in the study to classify bananitos maturity stages. SIMCA (Wold and Sjöström 1977) is a classifier based on PCA analysis. For each maturity stage in the calibration set, an individual PCA model was developed with selection of an optimal number of PCs by cross-validation (Djuris et al. 2013). The number of PCs in each PCA model can be different. When all three classes were modelled, a SIMCA classification model was generated by assembling all PCA models. The class membership was determined by projecting validation samples to each PCA model and calculating their orthogonal distances (Tominaga 1999). PLSDA (Brereton and Lloyd 2014) is a linear classification method based on partial least square (PLS) algorithm, where the dependant y-responses are replaced with the class category. PLSDA analysis provides an average percentage of classification error for all classes during calibration and cross-validation; this information was used for selecting the optimal number of latent variables (LVs). Unlike SIMCA and PLSDA, k-NN (Cover and Hart 1967) is a non-parametric classification method. The class membership of an unknown sample was determined by its k-nearest neighbours in the calibration set. The nearest neighbour to the unknown sample is the sample that has the smallest Euclidean distance. The class membership of an unknown sample is decided by the majority class that its k-nearest neighbours

Fig. 2 Spectral trimming and interpolation. **a** The 90 raw mean-spectra from bananito peel; **b** the data-points before and after interpolation



have. Therefore, k-NN classification result might be influenced by the choice of k.

Wavelength Selection Methods Two methods were applied in the study to select feather wavelengths to optimise the model. One method is based on regression vector (RV) and the other is based on variable importance in projection (VIP) scores. In a PLS-based model, the regression coefficients in the regression vector (**B**) indicates how much the spectral variables (**X**) will influence the measured responses (**Y**), as described below:

$$\mathbf{Y} = \mathbf{X}^* \mathbf{B} + \mathbf{E} \quad (4)$$

where **E** is the model residual.

For the RV method, a wavelength that has a large regression coefficient (absolute value) is regarded as significant to the measured responses **Y**, thus, this wavelength can be considered as an important wavelength.

For the VIP scores method, the VIP value of each wavelength specifies the importance (score) of each wavelength/variable in modelling for both **X** predictors and **Y** responses (Svante and Eriksson 2001). It can be calculated using the equation provided in the paper of Lu et al. (2014). The threshold value of the VIP score for important wavelengths selection is generally set at 1, since the mean of the squared VIP scores is equal to 1 (Chong and Jun 2005).

Classification Rule and Model Performance Cross-validation was performed using venetian blinds (7 splits) when modelling. To classify each sample in the prediction set, the probability of each sample being inside each class was calculated in the PLS_toolbox software. Classification rule used in the study was based on “class predicted most probable”, meaning that a sample was assigned to a class that has the largest probability value. In this case, each sample was assigned to one class.

Classification performance of three classifiers was evaluated by total correct classification rate (TCC%), it can be calculated by Eq. 5.

$$\text{TCC}(\%) = \frac{N_c}{N_{\text{total}}} \times 100 \quad (5)$$

where N_c represents the number of samples being correctly classified and N_{total} indicates the total number of samples in the prediction set.

Results and Discussion

Changes of Firmness, SSC and Colour

Table 2 shows the changes of bananito flesh firmness, SSC and peel colour that were related with ripeness stage. The mean value of flesh firmness for maturity stages 2, 4 and 6 fruits were 15.26 N, 10.81 N and 5.10 N, respectively, showing that flesh firmness decreased with an increase in maturity level. In contrast, soluble solids content of bananito flesh increased with increasing ripeness stage. The mean value of SSC in the flesh increased from 7.53 to 17.94% when fruit changed from stage 2 to stage 6. The loss of mechanical resistance and the increase in soluble solids content during ripening process is generally associated with the activity of ripening-related enzymes such as amylase, pectinase and cellulase. During fruit ripening, the primarily-stored starch is hydrolysed to small molecules such as sucrose, fructose and glucose, leading to a reduction in starch content and to an increase in total soluble solids (Garcia and Lajolo 1988). The pectinase present in the cell wall is responsible for the degradation of pectins, which, in turn, leads to a softer texture in banana fruits (Smith et al. 1990). The riper the fruit is, the softer the texture would be.

Table 2 Changes of firmness, SSC and colour parameters in bananito fruit of different maturity stages

		Maturity stage		
		2	4	6
Internal quality	Firmness (N)	15.16 ± 7.24 ^c	10.81 ± 6.56 ^b	5.10 ± 1.54 ^a
	SSC (%)	7.53 ± 2.35 ^a	9.88 ± 4.19 ^b	17.94 ± 1.17 ^c
Peel colour	<i>L</i> [*]	64.10 ± 1.16 ^a	65.37 ± 2.04 ^a	70.99 ± 4.39 ^b
	<i>a</i> [*]	− 13.68 ± 1.01 ^a	− 11.53 ± 3.79 ^b	3.05 ± 3.71 ^c
	<i>b</i> [*]	35.14 ± 1.17 ^a	36.21 ± 2.87 ^a	43.97 ± 4.20 ^b
	<i>Chroma</i>	37.79 ± 1.12 ^a	38.27 ± 2.56 ^a	44.28 ± 4.10 ^b
	<i>h</i> [°]	111.29 ± 1.61 ^c	107.70 ± 6.11 ^b	85.91 ± 5.03 ^a
	<i>YI</i>	43.90 ± 1.01 ^a	45.12 ± 2.52 ^a	54.08 ± 2.03 ^b

The data is presented as “mean ± standard deviation” for each maturity stage. Mean values in the same row with different subscript letters “a, b, c” are significantly different ($p < 0.05$) based on Bonferroni post hoc test

In terms of colour changes when fruit ripens, peel colour a^* (green-red) had a significant increase from maturity stage 2 to stage 6. For example, unripe bananitos at maturity stage 2 had a negative a^* value ($a^* = -13.68$), indicating the greenness of the peel colour; while ripe bananitos at maturity stage 6 showed a positive a^* value ($a^* = 3.05$), indicating the disappearance of greenness due to chlorophyll decomposition. The peel hue angle showed a significant and decreasing trend, with h° values changed from 111.3 (stage 2 fruit) to 85.9 (stage 6 fruit), revealing a change of the peel hue from green to yellow. It was noted that there was no statistical difference between stage 2 and stage 4 fruit in the peel colour parameters of L^* , b^* , *Chroma* and *YI*, however, the difference between fruits at stage 6 and the other maturity stages (2 and 4) was significantly different. The changing trend of the peel colour L^* and a^* showed by bananito fruit (*cv. Pisang mas*) in this study was in agreement with the findings of Sanaeifar et al. (2016) on conventional banana (*cv. Cavendish*). However, the a^* value of the Cavendish banana fruit ranged from -30 to 0 , which was lower than the a^* value of bananito fruit (ranged from -13.68 to 3.05).

Correlations Between Firmness, SSC and Colour Features

The correlation coefficients (r) between firmness, SSC and several peel colour parameters are shown in Table 3. When comparing the correlation coefficients of firmness or SSC to all colour parameters, it was observed that SSC had a higher r values, indicating that this maturity index has a closer connection with fruit colour. Sanaeifar et al. (2016) studied the correlation between peel colour (L^* , a^* and b^*) and SSC and firmness in bananas (*cv. Cavendish*), showing similar results. Among all correlation coefficients that were statistically significant, the highest r value was found between SSC and peel *YI* ($r = 0.92$), indicating the strongest relationship between them and the potential of using peel yellowness index to estimate SSC. Other peel colour parameters (L^* , a^* , b^* , *Chroma* and h°) had a slightly lower correlation to SSC when compared with peel *YI*,

Table 3 Correlation coefficients (r) between firmness, SSC and colour parameters

	Peel colour					
	L^*	a^*	b^*	<i>Chroma</i>	h°	<i>YI</i>
Firmness	-0.47	-0.58	-0.57	-0.51	0.61	-0.63
SSC	0.72	0.87	0.82	0.75	-0.89	0.92

Correlation coefficients shown in the above table are at significant level of $p < 0.001$

with r equal to 0.72, 0.87, 0.82, 0.75 and -0.89 , respectively.

Spectral Profiles of Bananito Fruit

Mean spectra (in common wavelength range of 400–740 nm) extracted from bananito fruits using spectrometer (Vis-point) and hyperspectral imaging (Vis-HSI) were investigated. Figure 3 shows the visible reflectance spectra collected from the bananito peel. When comparing raw peel spectra of Vis-point (Fig. 3a) and Vis-HSI (Fig. 3d), it was observed that the spectral profile showed a similar variation trend, and these spectral changes were in line with the study of conventional banana (Rajkumar et al. 2012). Take Vis-point spectra, for example the main spectral variation among the three groups of bananitos took place in the wavelength region of 540–740 nm. An obvious and relatively broad absorption band in bananito peel was found at around 675 nm, which is associated with chlorophyll (Li et al. 1997). Since spectroscopy is a quantitative analytical method, the reflectance intensity of spectral signal was related to the concentration of chlorophyll content presented in fruit skin. The degradation of chlorophyll in bananito peel during fruit maturation resulted in a pronounced reduction in spectral absorbance at 675 nm. As shown in Fig. 3 a, the reflectance intensity of bananito fruits of stage 2 was much smaller than that of stage 6 fruits, while the reflectance intensity of stage 4 fruits was in between stages 2 and 6. Dispersion in the Vis-point and Vis-HSI spectra was observed; however, the scattering effect in Vis-point spectra was larger than that in Vis-HSI. Since the mean spectra of Vis-point was obtained by averaging the spectra from six small areas on the peel, while the mean spectra of Vis-HSI was obtained by averaging all pixel spectra from whole fruit (except both ends), the scattering effect in Vis-HSI spectra might be averaged-out by including more variations from the sample. By use of SNV spectral pre-processing, undesired scattering effects were significantly reduced in the two sets of spectra, as displayed in Fig. 3 b and e.

PCA analysis was conducted for bananito peel spectra to explore and visualize their trends based on full wavelength spectra. With the application of SNV and mean-centring before PCA, the scores plots of Vis-point and Vis-HSI data were shown in Fig. 3 c and f, respectively. PCA scores plot shows the position of each bananito fruit in the determined PC1–PC2 space. Spectra that had similar spectral characteristics were close to each other. For Vis-point bananito peel data, the first PC (PC1) accounted for the greatest variance (95.44%) presented in the data, while the second PC (PC2) which was orthogonal to PC1 explained 3.76% of data variance. Similarly, for Vis-HSI data, the PC1 and PC2 accounted for 93.97% and 5.45% of data variance, respectively. In summary, the first two PCs in Vis-point and Vis-HSI captured over 99% of total variance in the data, demonstrating the capability of

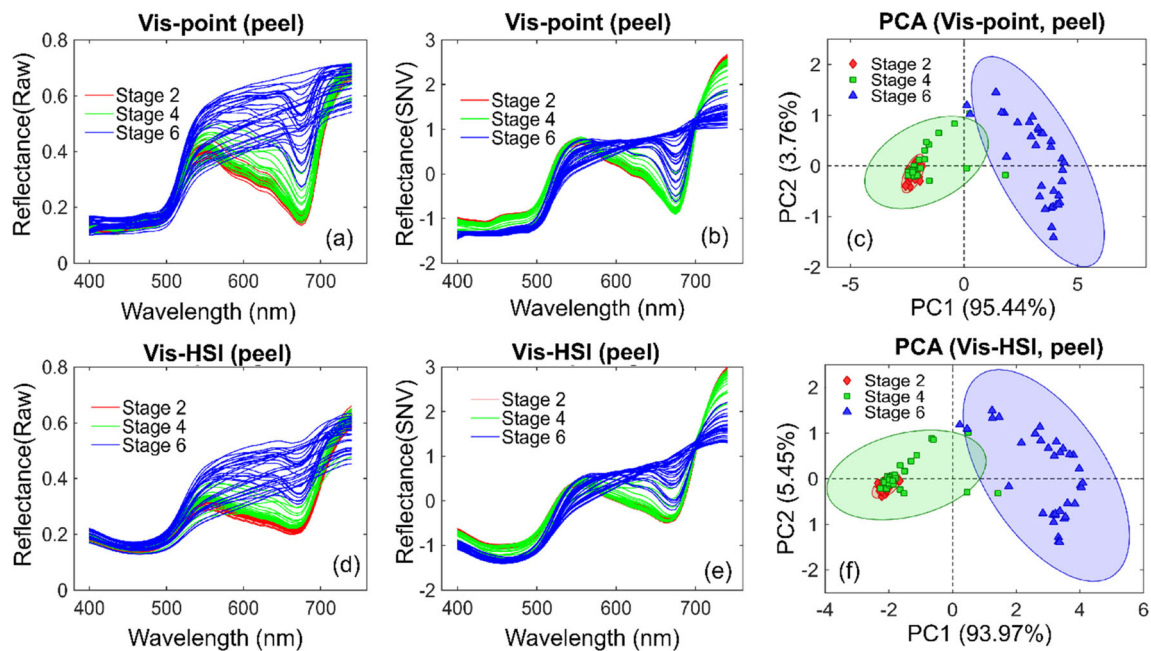


Fig. 3 Spectral profile and PCA analysis result of bananito fruit collected from spectrometer (Vis-point) and hyperspectral imaging (Vis-HSI)

PCA in dimension reduction. It was clearly shown in the scores plot that application of PCA yielded a distinct separation between bananito fruits in ripeness stage 2 and stage 6. However, a part of stage 4 samples was overlapped with stage 2 samples, which illustrated that the use of PCA was not sufficient for bananitos maturity stage classification and more complex supervised classification methods should be applied to improve classification accuracy.

Comparison of Classification Models

Three classification models, namely k-NN, SIMCA and PLSDA, were developed based on the full visible wavelengths of fruit samples in the calibration set (60 fruits in total). Cross validation was carried out in the calibration set using “venetian blinds” method that splits the data into 7 subsets. Selections of the optimal k value in k-NN, the

Table 4 Ripeness classification result in the prediction set based on three classification models (TCC total correct classification)

Model	Data	Confusion matrix in prediction set			No. of TCC samples	
			Predicted stage 2	Predicted stage 4		Predicted stage 6
k-NN	Vis-point ($k = 4$)	Stage 2	9	1	0	25
		Stage 4	3	6	0	
		Stage 6	0	1	10	
k-NN	Vis-HSI ($k = 3$)	Stage 2	7	3	0	24
		Stage 4	3	6	0	
		Stage 6	0	0	11	
SIMCA	Vis-point (PCs: 4, 3, 3)	Stage 2	8	2	0	24
		Stage 4	3	6	0	
		Stage 6	0	1	10	
SIMCA	Vis-HSI (PCs: 4, 3, 3)	Stage 2	10	0	0	26
		Stage 4	4	5	0	
		Stage 6	0	0	11	
PLSDA	Vis-point (LV = 4)	Stage 2	7	2	0	24
		Stage 4	3	6	0	
		Stage 6	0	0	11	
PLSDA	Vis-HSI (LV = 5)	Stage 2	9	1	0	28
		Stage 4	1	8	0	
		Stage 6	0	0	11	

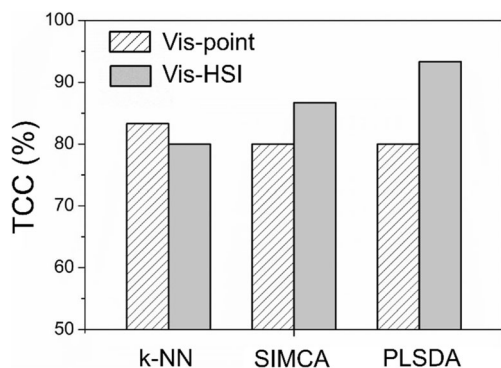


Fig. 4 Comparison of classification accuracy of three classifiers using Vis-point and Vis-HSI data of bananito

number of PCs of each PCA model in SIMCA and the number of LVs in PLSDA were based on the average-classification error of three classes in the cross validation. The results were shown in the column 2 of Table 4. For example, by selecting 4 LVs in the PLSDA modelling, the average classification error of the three ripeness classes was minimal. Once a calibration model was developed, an external set of samples (a test set or a prediction set) was applied to model to validate the model performance. Table 4 presents the ripeness classification details of a prediction set using bananito peel spectra. It can be seen from the confusion matrix that classification between stage 2 and stage 6 fruits was identical; no fruits in ripeness stage 2 or 6 were misclassified into stage 6 or 2. However, only few bananito fruits in stage 2 or 6 were misclassified to its adjacent neighbour (stage 4). For example, of the actual ten samples in stage 2, k-NN model using Vis-point peel spectra predicted that nine samples were in stage 2 and one sample was misclassified as stage 4. Of the actual eleven samples in stage 6, the above model predicted that ten samples were in stage 6 and one sample was in stage 4. However, all classification models in Table 4 had difficulties in classifying bananito fruits of maturity stage 4. Using the peel visible spectra, some bananitos in stage 4 were misclassified to stage 2. This result could be explained by

the similarity of peel spectra between stages 2 and 4, as shown in Fig. 3.

From the confusion matrix shown in Table 4, the total correct classification rate (TCC%) for each classification model can be calculated. Figure 4 shows the TCC% result of each model for bananito ripeness classification. It is found that the TCC% of Vis-HSI was generally higher than that of Vis-point. That is to say, the use of hyperspectral imaging provides higher classification results than the use of spectrometer. A possible explanation for this result would be the fact that spectrometer only covered some small areas from the sample whereas hyperspectral imaging system could capture spectral information from the whole sample. Since the peel colour of bananito fruits was not uniform from stem to tip (Chen and Ramaswamy 2002), the mean-spectrum extracted from HSI was more representative for ripeness classification. The highest classification accuracy was achieved by using Vis-HSI peel spectra in the PLSDA model, giving a TCC% of 93.3%. Based on the same data, the SIMCA model ranked in the second place for bananito maturity classification, yielding a TCC% of 86.7%.

Optimal Wavelengths Selection

Feature wavelengths selection is an important strategy to investigate optical properties of bananitos in relation to its ripeness stage. It is beneficial to simplify modelling process and to improve model accuracy by reducing data inputs and removing some unrelated variables (Liu et al. 2014). Since the PLSDA model established on peel Vis-HSI spectra (full wavelengths) gave the highest ripeness prediction accuracy, selection of optimal wavelength was conducted on peel Vis-HSI spectra using PLSDA. Two approaches, namely regression vector (RV) and variable importance in projection (VIP) scores, as detailed in “Wavelength Selection Methods”, were applied to identify key wavelengths for ripeness classification. In the regression vector plot derived from the three PLSDA classification models for the three maturity stages, wavelengths that have the largest absolute value of regression

Fig. 5 Selection of feature wavelengths based on regression vector (a) or VIP scores (b) using peel Vis-HSI data

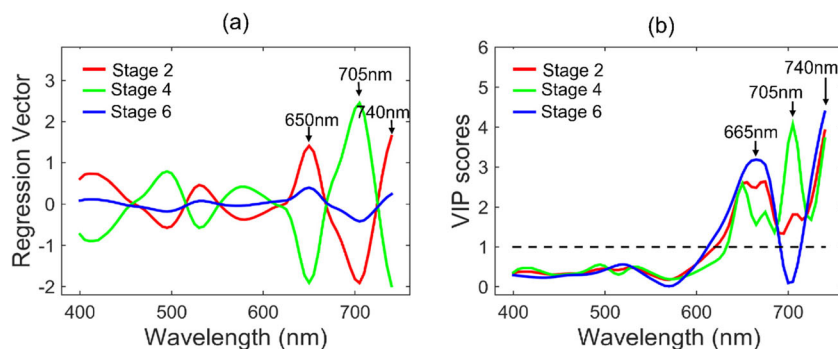


Table 5 Classification results of full-wavelength PLSDA model and simplified PLSDA model

Model	Wavelengths (nm)	Wavelength selection method	No. of latent variable	No. of misclassified samples	TCC (%)
PLSDA-1	400–740	–	5	2	93.3
PLSDA-2	650, 705, 740	RV	2	2	93.3
PLSDA-3	665, 705, 740	VIP	3	2	93.3

coefficients indicate its great contribution to that model. In the VIP scores plot of the PLSDA classification model, wavelengths of VIP scores over 1 are generally considered to be important wavelengths for the model (Pu and Sun 2016). Based on these selection rules, three feature wavelengths (650 nm, 705 nm and 740 nm) were identified from the regression vector plot (as shown in Fig. 5a) and three feature wavelengths (665 nm, 705 nm and 740 nm) were identified in the VIP scores plot (as shown in Fig. 5b).

The wavelengths 650 nm and 665 nm are in the range of secondary absorption peaks of the reaction centre of the photosystem II (of the photosynthesis), and of the “light-harvesting” *chlorophyll a* and *b* protein complexes (650–680 nm) (Machlis and Torrey 1956). The wavelength 705 nm belongs to the so called “red edge”, a spectral range that has a very high sensitivity to changes in *Chlorophyll a* and *a+b*. This “red edge” occurs between the wavelengths 680 nm and 710 nm, with a maximum sensitivity at 700 nm (Filella and Penuelas 1994). A close relationship between the “red edge” and the leaf chlorophyll concentration has been demonstrated (Lichtenthaler et al. 1996). The wavelength 740 nm is a minor absorption feature for water (Cubeddu et al. 2003).

Based on the peel Vis-HSI spectra that were processed with SNV and mean-centring, spectral intensities at the three important wavelengths selected by RV (650 nm, 705 nm and 740 nm) or VIP (665 nm, 705 nm and 740 nm) were extracted to construct a simplified PLSDA model. Classification results of full-wavelength PLSDA model (PLSDA-1) and the two simplified PLSDA models (PLSDA-2 and PLSDA-3) were compared and shown in Table 5. Comparing PLSDA-2 and PLSDA-3, both models yielded an equivalent classification

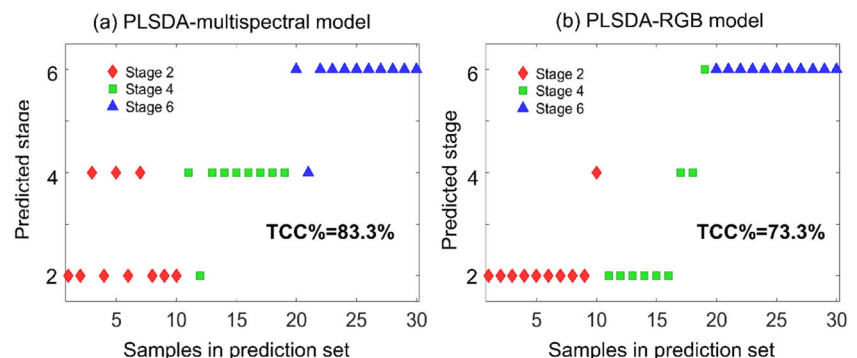
result. The classification accuracy of PLSDA-2 or PLSDA-3 was the same as PLSDA-1, indicating the capability of using selected wavelengths to classify bananitos’ maturity stage. The best classification model in the study was PLSDA-2, as less number of latent variables was used.

Multispectral Imaging

Though a high classification accuracy (TCC% = 93.3%) was obtained by the simplified PLSDA-2 or PLSDA-3 model, it should be noted that these reduced wavelength models were based on spectral values extracted after SNV and mean-centring pre-treatments were applied to the full spectrum. This means that the full wavelengths are still needed to achieve such a good classification performance.

From the perspective of industrial application, it would be more practical to develop a multispectral imaging system equipped with few wavelength channels for bananitos’ ripeness classification. To investigate the potential of using feature wavelengths (650 nm, 705 nm and 740 nm) identified in the study to classify bananitos’ maturity stage, the raw spectral data (the peel Vis-HSI spectra) at the above-mentioned three wavelengths were extracted. These raw spectral data were processed with SNV and mean-centring, then a PLSDA model (named PLSDA-multispectral) was established. Figure 6 a shows the corresponding classification result for the validation set. Five samples were misclassified (three samples in stage 2 were misclassified as stage 4, one sample in stage 4 was misclassified as stage 2, and one sample in stage 6 was misclassified as stage 4), yielding a total correct classification rate of 83.3% for the PLSDA-multispectral model.

Fig. 6 Comparison of simplified classification models. **a** Classification result of PLSDA-multispectral model developed using the raw spectral data extracted from the three feature wavelengths 650 nm, 705 nm and 740 nm; **b** classification result of PLSDA-RGB model developed based on the raw spectral data extracted from three wavelengths 650 nm (red), 500 nm (green) and 450 nm (blue)



Currently, commercial RGB cameras are available for fruit ripeness classification based on the colour images. It is interesting to compare the classification results obtained by the PLSDA-multispectral model with the results obtained by the RGB camera. Here in the study, an RGB spectral dataset was constructed by selecting three wavelengths (650 nm, 500 nm and 450 nm) as the Red, Green and Blue channel from the hyperspectral data (Cheng et al. 2016). A PLSDA model (named PLSDA-RGB) was then developed for bananito maturity classification, the classification result is shown in Fig. 6 b. In total, eight samples were misclassified (one sample in stage 2 were misclassified as stage 4, six samples in stage 4 were misclassified as stage 2 and one sample in stage 4 was misclassified as stage 6), giving a total correct classification rate of 73.3% for the PLSDA-RGB model.

When comparing Fig. 6 a to Fig. 6 b, it can be seen that the PLSDA-multispectral model (TCC% = 83.3%) performs better than the PLSDA-RGB model (TCC% = 73.3%), indicating the feasibility and potential of developing a multispectral imaging system with the use of only three wavebands (650 nm, 705 nm and 740 nm) identified in the study for bananito ripeness classification.

Conclusions

This study investigated variations of firmness, SSC and colour of bananito during the ripening process, and the use of spectrometer and hyperspectral imaging for ripeness stage classification. Results showed that fruit firmness decreased whereas SSC increased during fruit ripening. SSC had a closer linkage to peel colour parameters as compared with firmness. The peel yellowness index was found to have the strongest correlation to SSC, with correlation coefficient of $r = 0.92$ obtained. A PLSDA model developed on full wavelength Vis-HSI spectral data performed the best result (TCC = 93.3%) in bananito ripeness classification. The three selected wavelengths (650 nm, 705 nm and 740 nm) obtained by RV method gave a comparative classification result to full wavelengths, and worked better than the RGB wavelengths. By comparing spectrometer and hyperspectral imaging, this study suggested the use of visible hyperspectral imaging technique for non-destructive and accurate classification of bananito fruits. Potential of using the three feature wavebands to construct a multispectral imaging system for industrial application has also been demonstrated.

Acknowledgments The authors would like to thank Annamaria Stellari for technical support and the AL.MA s.r.l. company in Milan, Italy, for bananito fruits supply. UCD-CSC Scholarship Scheme supported by University College Dublin (UCD) and China Scholarship Council (CSC) was also acknowledged for this study.

Compliance with Ethical Standards

Conflict of Interest Yuan-Yuan Pu declares that she has no conflict of interest. Da-Wen Sun declares that he has no conflict of interest. Marina Buccheri declares that she has no conflict of interest. Maurizio Grassi declares that he has no conflict of interest. Tiziana M.P. Cattaneo declares that she has no conflict of interest. Aoife Gowen declares that she has no conflict of interest.

Ethical Approval This article does not contain any studies with human participants performed by any of the authors.

Informed Consent A statement regarding informed consent is not applicable for this study.

References

- ASTM 1988. Standard Test Method for Yellowness Index of Plastics (ASTM D 1925-70)
- Blankenship SM, Ellsworth DD, Powell RG (1993) A ripening Index for banana fruit based on starch content. *HortTechnology* 3:338–339
- Brereton RG, Lloyd GR (2014) Partial least squares discriminant analysis: taking the magic away. *J Chemom* 28:213–225
- Chen CR, Ramaswamy HS (2002) Color and texture change kinetics in ripening bananas. *LWT Food Sci Technol* 35:415–419
- Cheng JH, Sun DW, Nagata M, Tallada JG (2016) Chapter 13 - Quality Evaluation of Strawberry. In: SUN D-W (ed) *Computer Vision Technology for Food Quality Evaluation (Second Edition)*. Academic Press, San Diego
- Cheng W, Sun D-W, Cheng J-H (2016a) Pork biogenic amine index (BAI) determination based on chemometric analysis of hyperspectral imaging data. *LWT Food Sci Technol* 73:13–19
- Cheng W, Sun D-W, Pu H, Liu Y (2016b) Integration of spectral and textural data for enhancing hyperspectral prediction of K value in pork meat. *LWT Food Sci Technol* 72:322–329
- Cheng W, Sun D-W, Pu H, Wei Q (2017) Chemical spoilage extent traceability of two kinds of processed pork meats using one multispectral system developed by hyperspectral imaging combined with effective variable selection methods. *Food Chem* 221:1989–1996
- Cheng W, Sun D-W, Pu H, Wei Q (2018) Heterospectral two-dimensional correlation analysis with near-infrared hyperspectral imaging for monitoring oxidative damage of pork myofibrils during frozen storage. *Food Chem* 248:119–127
- Chong I-G, Jun C-H (2005) Performance of some variable selection methods when multicollinearity is present. *Chemom Intell Lab Syst* 78:103–112
- Cortés V, Blasco J, Aleixos N, Cubero S, Talens P (2017) Visible and Near-Infrared Diffuse Reflectance Spectroscopy for Fast Qualitative and Quantitative Assessment of Nectarine Quality. *Food Bioprocess Technol* 10:1755–1766
- Cortés V, Cubero S, Blasco J, Cubero S, Talens P (2019) In-line Application of Visible and Near-Infrared Diffuse Reflectance Spectroscopy to Identify Apple Varieties. *Food Bioprocess Technol*. <https://doi.org/10.1007/s11947-019-02268-0>
- Cover T, Hart P (1967) Nearest neighbor pattern classification. *IEEE Trans Inf Theory* 13:21–27
- Cubeddu R, Pifferi A, Taroni P, Torricelli A (2003) Chapter 14 Spectroscopic techniques for analysing raw material quality. In: TOTHILL I (ed) *Rapid and On-Line Instrumentation for Food Quality Assurance*. Elsevier Science, Amsterdam
- Dai Q, Cheng J-H, Sun D-W, Zhu Z, Pu H (2016) Prediction of total volatile basic nitrogen contents using wavelet features from

- visible/near-infrared hyperspectral images of prawn (*Metapenaeus ensis*). *Food Chem* 197:257–265
- Davey MW, Saeys W, Hof E, Ramon H, Swennen RL, Keulemans J (2009) Application of visible and near-infrared reflectance spectroscopy (Vis/NIRS) to determine carotenoid contents in banana (*Musa spp.*) fruit pulp. *J Agric Food Chem* 57:1742–1751
- Djuris J, Ibric S, Djuric Z (2013) Chapter 4: Chemometric methods application in pharmaceutical products processes analysis control. In: Djuris J (ed) *Computer-aided applications in pharmaceutical technology*. Elsevier Science, Amsterdam
- Du CJ, Sun D-W (2005) Pizza sauce spread classification using colour vision and support vector machines. *J Food Eng* 66:137–145
- Eduard L, Evor LH, Julian WG, Stefano F (1999) Non-destructive banana ripeness determination using a neural network-based electronic nose. *Meas Sci Technol* 10:538
- Filella I, Penuelas J (1994) The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *Int J Remote Sens* 15:1459–1470
- FRESHPLAZA. (2018) *The Bananito is a very popular Spreafico product* [Online]. Available: <https://www.freshplaza.com/article/191921/The-Bananito-is-a-very-popular-Spreafico-product/> [Accessed 08 January 2019]
- García E, Lajolo FM (1988) Starch Transformation during banana ripening: the amylase and glucosidase behavior. *J Food Sci* 53:1181–1186
- Gowen AA, Taghizadeh M, O'donnell CP (2009) Identification of mushrooms subjected to freeze damage using hyperspectral imaging. *J Food Eng* 93:7–12
- Hashim N, Janius R, Baranyai L, Rahman R, Osman A, Zude M (2012) Kinetic model for colour changes in bananas during the appearance of chilling injury symptoms. *Food Bioprocess Technol* 5:2952–2963
- Hashim N, Pflanz M, Regen C, Janius RB, Abdul Rahman R, Osman A, Shitan M, Zude M (2013) An approach for monitoring the chilling injury appearance in bananas by means of backscattering imaging. *J Food Eng* 116:28–36
- He H-J, Sun D-W (2015) Microbial evaluation of raw and processed food products by Visible/Infrared, Raman and Fluorescence spectroscopy. *Trends Food Sci Technol* 46:199–210
- Jackman P, Sun D-W, Allen P (2011) Recent advances in the use of computer vision technology in the quality assessment of fresh meats. *Trends Food Sci Technol* 22:185–197
- Jaiswal P, Jha SN, Kaur PP, Bhardwaj R, Singh AK, Wadhawan V (2014) Prediction of textural attributes using color values of banana (*Musa sapientum*) during ripening. *J Food Sci Technol* 51:1179–1184
- Kajuna STAR, Bilanski WK, Mittal GS (1998) Color changes in bananas and plantains during storage. *J Food Process Preserv* 22:27–40
- Li M, Slaughter DC, Thompson JF (1997) Optical chlorophyll sensing system for banana ripening. *Postharvest Biol Technol* 12:273–283
- Li J, Ould Eleya MM, Gunasekaran S (2006) Gelation of whey protein and xanthan mixture: effect of heating rate on rheological properties. *Food Hydrocoll* 20:678–686
- Li W, Shao Y, Chen W, Jia W (2011) The effects of harvest maturity on storage quality and sucrose-metabolizing enzymes during banana ripening. *Food Bioprocess Technol* 4:1273–1280
- Lichtenthaler HK, Gitelson A, Lang M (1996) Non-destructive determination of chlorophyll content of leaves of a green and an aurea mutant of tobacco by reflectance measurements. *J Plant Physiol* 148:483–493
- Liew CY, Lau CY (2012) Determination of quality parameters in Cavendish banana during ripening by NIR spectroscopy. *Int Food Res J* 19:751–758
- Liu D, Sun D-W, Zeng X-A (2014) Recent advances in wavelength selection techniques for hyperspectral image processing in the food industry. *Food Bioprocess Technol* 7:307–323
- Liu F, He Y, Wang L, Sun G (2011) Detection of Organic Acids and pH of Fruit Vinegars Using Near-Infrared Spectroscopy and Multivariate Calibration. *Food Bioprocess Technol* 4:1331–1340
- Lu B, Castillo I, Chiang L, Edgar TF (2014) Industrial PLS model variable selection using moving window variable importance in projection. *Chemom Intell Lab Syst* 135:90–109
- Liu Y, Pu H, Sun D-W (2017) Hyperspectral imaging technique for evaluating food quality and safety during various processes: a review of recent applications. *Trends Food Sci Technol* 69:25–35
- Liu Y, Sun D-W, Cheng J-H, Han Z (2018) Hyperspectral imaging sensing of changes in moisture content and color of beef during microwave heating process. *Food Anal Methods* 11:2472–2484
- Ma J, Sun D-W, Pu H (2017) Model improvement for predicting moisture content (MC) in pork longissimus dorsi muscles under diverse processing conditions. by hyperspectral imaging. *J Food Eng* 196:65–72
- Ma J, Pu H, Sun D-W (2018) Predicting intramuscular fat content variations in boiled pork muscles by hyperspectral imaging using a novel spectral pre-processing technique. *LWT Food Sci Technol* 94:119–128
- Machlis, L. & Torrey, J. G. (1956) *Plants in action. A laboratory manual of plant physiology. Plants in action. A laboratory manual of plant physiology*
- Magwaza L. S, Opara U. L, Nieuwoudt H, Cronje P. J. R, Saeys W, Nicolai B (2012) NIR Spectroscopy Applications for Internal and External Quality Analysis of Citrus Fruit—A Review. *Food Bioprocess Technol* 5:425–444
- Mclaren K (1976) XIII—The development of the CIE 1976 ($L^* a^* b^*$) uniform colour space and colour-difference formula. *J Soc Dye Colour* 92:338–341
- Mendoza F, Aguilera JM (2004) Application of image analysis for classification of ripening bananas. *J Food Sci* 69:E471–E477
- Mohapatra D, Mishra S, Singh CB, Jayas DS (2011) Post-harvest processing of banana: opportunities and challenges. *Food Bioprocess Technol* 4:327–339
- Morsy N, Sun D-W (2013) Robust linear and non-linear models of NIR spectroscopy for detection and quantification of adulterants in fresh and frozen-thawed minced beef. *Meat Sci* 93:292–302
- Pan Y, Sun D-W, Cheng J-H, Han Z (2018) Non-destructive detection and screening of non-uniformity in microwave sterilization using hyperspectral imaging analysis. *Food Anal Methods* 11:1568–1580
- Pathare P. B, Opara U. L, Al-said F. A.-J (2013) Colour Measurement and Analysis in Fresh and Processed Foods: A Review. *Food and Bioprocess Technol* 6:36–60
- Pu Y-Y, Sun D-W (2016) Prediction of moisture content uniformity of microwave-vacuum dried mangoes as affected by different shapes using NIR hyperspectral imaging. *Innovative Food Sci Emerg Technol* 33:348–356
- Pu Y-Y, Feng Y-Z, Sun D-W (2015) Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: a review. *Compr Rev Food Sci Food Saf* 14:176–188
- Quevedo R, Mendoza F, Aguilera JM, Chanona J, Gutiérrez-López G (2008) Determination of senescent spotting in banana (*Musa cavendish*) using fractal texture Fourier image. *J Food Eng* 84:509–515
- Rajkumar P, Wang N, Eimasry G, Raghavan GSV, Gariepy Y (2012) Studies on banana fruit quality and maturity stages using hyperspectral imaging. *J Food Eng* 108:194–200
- Rambo MKD, Ferreira MMC, Amorim EP (2016) Multi-product calibration models using NIR spectroscopy. *Chemom Intell Lab Syst* 151:108–114
- Ravikanth L, Jayas DS, White NDG, Fields PG, Sun D-W (2017) Extraction of Spectral Information from Hyperspectral Data and Application of Hyperspectral Imaging for Food and Agricultural Products. *Food Bioprocess Technol* 10:1–33
- Robinson JC, Saúco VG (2010) *Bananas and Plantains*. CABI, Wallingford

- Sanaeifar A, Bakhshipour A, De IA Guardia M (2016) Prediction of banana quality indices from color features using support vector regression. *Talanta* 148:54–61
- Smith NJ, Seymour GB, Jeger MJ, Tucker GA (1990) Cell wall changes in bananas and plantains. In: International Society for Horticultural Science (ISHS). Leuven, Belgium, pp 283–290
- Steffens C, Franceschi E, Corazza FC, Herrmann PSP Jr, Oliveira JV (2010) Gas sensors development using supercritical fluid technology to detect the ripeness of bananas. *J Food Eng* 101:365–369
- Sun D-W, Brosnan T (2003) Pizza quality evaluation using computer vision - Part 2 - Pizza topping analysis. *J Food Eng* 57:91–95
- Surya Prabha D, Satheesh Kumar J (2015) Assessment of banana fruit maturity by image processing technique. *J Food Sci Technol* 52:1316–1327
- Svante W, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst* 58:109–130
- Tominaga Y (1999) Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs, and k-NN. *Chemom Intell Lab Syst* 49:105–115
- USDA. 2001. *Banana ripening guide* [Online]. Available: https://www.ams.usda.gov/sites/default/files/media/Bananas_Visual_Aid%5B1%5D.pdf [Accessed March 2017]
- Wang HH, Sun D-W (2003) Assessment of cheese browning affected by baking conditions using computer vision. *J Food Eng* 56:339–345
- Wang N-N, Sun D-W, Yang Y-C, Pu H, Zhu Z (2015a) Recent advances in the application of hyperspectral imaging for evaluating fruit quality. *Food Anal Methods*:1–14
- Wang N-N, Yang Y-C, Sun D-W, Pu H, Zhu Z (2015b) Shelf-life prediction of ‘Gros Michel’ bananas with different browning levels using hyperspectral reflectance imaging. *Food Anal Methods* 8:1173–1184
- Wang L, Sun D-W, Pu H, Cheng J-H (2017a) Quality analysis, classification, and authentication of liquid foods by near-infrared spectroscopy: a review of recent research developments. *Crit Rev Food Sci Nutr* 57:1524–1538
- Wang K, Sun D-W, Pu H, Wei Q (2017b) Principles and applications of spectroscopic techniques for evaluating food protein conformational changes: a review. *Trends Food Sci Technol* 67:207–219
- Wold, S. & Sjöström, M. 1977. SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy. *Chemometrics: Theory and Application*. AMERICAN CHEMICAL SOCIETY
- Xu J-L, Riccioli C, Sun D-W (2015) An overview on nondestructive spectroscopic techniques for lipid and lipid oxidation analysis in fish and fish products. *Compr Rev Food Sci Food Saf* 14:466–477
- Zheng C, Sun D-W, Zheng L (2006) Correlating colour to moisture content of large cooked beef joints by computer vision. *J Food Eng* 77:858–863
- Zude M (2003) Non-destructive prediction of banana fruit quality using VIS/NIR spectroscopy. *Fruits* 58:135–142

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.